Econometric Theory, **29**, 2013, 609–628.
doi:10.1017/S0266466612000485

A FIXED-*b* PERSPECTIVE ON THE PHILLIPS–PERRON UNIT ROOT TESTS

TIMOTHY J. VOGELSANG
Michigan State University

MARTIN WAGNER
Institute for Advanced Studies, Vienna
and
Frisch Centre for Economic Research, Oslo

In this paper we extend fixed-*b* asymptotic theory to the nonparametric Phillips–Perron (PP) unit root tests. We show that the fixed-*b* limits depend on nuisance parameters in a complicated way. These nonpivotal limits provide an alternative theoretical explanation for the well-known finite-sample problems of the PP tests. We also show that the fixed-*b* limits depend on whether deterministic trends are removed using one-step or two-step detrending approaches. This is in contrast to the asymptotic equivalence of the one- and two-step approaches under a consistency approximation for the long-run variance estimator. Based on these results we introduce modified PP tests that allow for asymptotically pivotal fixed-*b* inference. The theoretical analysis is cast in the framework of near-integrated processes, which allows us to study the asymptotic behavior both under the unit root null hypothesis and for local alternatives. The performance of the original and modified PP tests is compared by means of local asymptotic power and a small finite-sample simulation study.

1. INTRODUCTION

In this paper we extend the fixed-*b* asymptotic theory of Kiefer and Vogelsang (2005) to the well-known unit root tests of Phillips and Perron (1988), i.e., the PP tests. We focus on the case where the PP tests are constructed using nonparametric kernel estimators of the long-run variance. We find that the fixed-*b* limits of the PP tests are not pivotal and furthermore also depend on whether deterministic trends are removed using one-step or two-step detrending methods. Our results are in contrast to existing results based on consistency of the long-run variance estimator

The authors thank a co-editor (A.M. Robert Taylor), two referees, conference participants at the DAGStat 2010 meeting in Dortmund, the 5th CSDA International Conference on Computational and Financial Econometrics in London, and seminar participants at the Institute for Advanced Studies in Vienna for helpful comments and suggestions. The authors acknowledge financial support from the Jubiläumsfonds of the Oesterreichische Nationalbank (grant 13398). The usual disclaimer applies. Address correspondence to Timothy J. Vogelsang, Dept. of Economics, 486 W. Circle Drive, 110 Marshall-Adams Hall, East Lansing, MI 48224-1038; e-mail: tjv@msu.edu.

in which case the asymptotic distributions of the PP tests are pivotal and are the same for one- and two-step detrending for the deterministic polynomial trends considered (see also Remark 1 in Section 2.3). Our finding of a nonpivotal fixed-b limit provides an alternative explanation for the often inadequate finite-sample performance of the PP tests (Perron and Ng, 1996; Schwert, 1989).

We propose a simple adjustment to the PP tests that provides a pivotal fixed-b limit under the unit root null. The theoretical analysis is performed using the framework of near-integrated process (cf. Phillips, 1987), which allows the derivation of limiting distributions both under the unit root null hypothesis and under local alternatives (to study local asymptotic power).

The remainder of the paper is organized as follows. In the next section we provide the fixed-b limits of the one- and two-step detrended versions of the PP tests. In Section 3 we propose the adjustment that restores an asymptotically pivotal fixed-b limit. Section 4 provides some limited finite-sample results and Section 5 briefly summarizes and concludes. All proofs are relegated to the Appendix. Additional material available upon request provides fixed-b critical values for five kernel functions (Bartlett, Bohman, Daniell, Parzen, and quadratic spectral (QS)), for the specifications without deterministic components, with intercept only, and with intercept and linear trend. For the latter two specifications of the deterministic component, critical values are available for both one- and two-step detrending. MATLAB programs to compute the modified test statistics and to perform inference using the fixed-b critical values are also available.

2. THE FIXED-*b* LIMITS OF THE PHILLIPS–PERRON TESTS

We assume that the data are generated according to

$$y_t = D_t'\theta + y_t^0, \qquad t = 1, \ldots, T, \tag{1}$$

$$y_t^0 = \left(1 - \frac{c}{T}\right) y_{t-1}^0 + u_t, \tag{2}$$

where, when deterministic components are included, $D_t := [1, t, t^2, \ldots, t^q]'$ for some $0 \leq q < \infty$. When $c = 0$, y_t^0 is a unit root process, and values of $c > 0$ correspond to near-integrated (in the terminology of Phillips, 1987) stationary (for fixed T) alternatives.

The key assumption is that the process u_t satisfies a functional central limit theorem (FCLT), i.e., for $T \to \infty$ it holds that

$$T^{-1/2} \sum_{t=1}^{[rT]} u_t \Rightarrow \omega W(r), \tag{3}$$

where $[rT]$ is the integer part of rT with $r \in [0, 1]$, $\Rightarrow$ signifies weak convergence, $W(r)$ is a standard Wiener process, and $0 < \omega^2 < \infty$ is the long-run variance of u_t.

Assuming for notational simplicity that u_t is covariance stationary with summable autocovariance function, with $\gamma_j := \mathbb{E}(u_t u_{t-j})$, we have

$$\omega^2 := \gamma_0 + 2\sum_{j=1}^{\infty} \gamma_j. \tag{4}$$

As is common, we use σ^2 to denote γ_0, and we furthermore define the half long-run variance $\lambda := \frac{1}{2}\left(\omega^2 - \sigma^2\right)$. In case u_t is not assumed to be stationary, ω^2 and σ^2 are defined as $\omega^2 := \lim_{T\to\infty} \mathbb{E}\left(\frac{1}{T}\left(\sum_{t=1}^{T} u_t\right)^2\right)$ and $\sigma^2 := \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}(u_t^2)$ with these limits assumed to exist; compare Phillips and Perron (1988). Sufficient conditions for the FCLT (3) can be found in Phillips and Solo (1992), or more specifically also in Phillips and Perron (1988) and in Sims, Stock, and Watson (1990), who consider similar deterministic components as we do. It is well known (see Phillips, 1987) that given (2) and (3) it follows that for $T \to \infty$

$$T^{-1/2} y^0_{[rT]} \Rightarrow \omega V_c(r),$$

where $V_c(r) := \int_0^r e^{-c(r-s)} dW(s)$.

Before we can turn to the analysis of the tests' asymptotic behavior, we need to define several additional quantities. Define $D(r) := [1, \ldots, r^q]'$ and the correspondingly detrended process, $\widetilde{V}_c(r)$, and generalized Brownian bridge, $\widehat{W}(r)$, as

$$\widetilde{V}_c(r) := V_c(r) - D(r)'\left(\int_0^1 D(s)D(s)'ds\right)^{-1}\int_0^1 D(s)V_c(s)ds, \tag{5}$$

$$\widehat{W}(r) := W(r) - \int_0^r D(s)'ds\left(\int_0^1 D(s)D(s)'ds\right)^{-1}\int_0^1 D(s)dW(s). \tag{6}$$

A slight variant of $\widetilde{V}_c(r)$ is also needed and is defined as

$$\dot{\widetilde{V}}_c(r) := W(r) - \int_0^r \left(\dot{D}(s) + cD(s)\right)' ds\left(\int_0^1 D(s)D(s)'ds\right)^{-1}\int_0^1 D(s)V_c(s)ds,$$

where $\dot{D}(r) := \partial D(r)/\partial r = \left[0, 1, 2r, \ldots, qr^{q-1}\right]'$. Note that $\int_0^r \dot{D}(s)ds = \left[0, r, r^2, \ldots, r^q\right]'$ is simply $D(r)$ with its first element replaced with 0. For the pure unit root case, $c = 0$, because $V_0(r) = W(r)$, it follows that $\widetilde{V}_0(r)$ and $\dot{\widetilde{V}}_0(r)$ are similar but different stochastic processes. As we shall see subsequently, this difference is the effect of using either one- or two-step detrending.

In the rest of the paper we will use throughout a subscript $i = 1, 2$ to refer to quantities related to either one- or two-step detrending, e.g., $\widetilde{y}_{t,1}$ refers to the one-step detrended y_t.

2.1. One-Step Approach

The one-step approach is based on estimating the regression model

$$y_t = D_t'\delta + \alpha y_{t-1} + u_t, \qquad t = 2, \ldots, T,$$

with the null hypothesis of interest being $H_0 : \alpha = 1$. Using the Frisch–Waugh theorem, the deterministic components can be eliminated, and one can focus on the regression

$$\widetilde{y}_{t,1} = \alpha \widetilde{y}_{t-1,1} + \widetilde{u}_t, \qquad t = 2, \ldots, T, \tag{7}$$

with $\widetilde{y}_{t,1} := y_t - D_t'\left(D_T' D_T\right)^{-1} D_T' Y_T$, $\widetilde{y}_{t-1,1} := y_{t-1} - D_t'\left(D_T' D_T\right)^{-1} D_T' Y_{T-1}$, and $\widetilde{u}_t := u_t - D_t'\left(D_T' D_T\right)^{-1} D_T' U_T$, using the notation $Y_T := [y_2, \ldots, y_T]'$, $Y_{T-1} := [y_1, \ldots, y_{T-1}]'$, $U_T := [u_2, \ldots, u_T]'$, and $D_T := [D_2, \ldots, D_T]'$.

The one-step PP unit root tests are based on the ordinary least squares (OLS) estimator $\widehat{\alpha}_1 := \left(\sum_{t=2}^T \widetilde{y}_{t,1}\widetilde{y}_{t-1,1}\right)/\left(\sum_{t=2}^T \widetilde{y}_{t-1,1}^2\right)$ of α from (7), respectively, the t-statistic

$$t_{\alpha_1} := \frac{\widehat{\alpha}_1 - 1}{\sqrt{\widehat{\sigma}_1^2 \left(\sum_{t=2}^T \widetilde{y}_{t-1,1}^2\right)}},$$

with $\widehat{\sigma}_1^2 := \frac{1}{T}\sum_{t=2}^T \widehat{u}_{t,1}^2$ and $\widehat{u}_{t,1} := \widetilde{y}_{t,1} - \widehat{\alpha}_1 \widetilde{y}_{t-1,1}$. Furthermore, denote the estimated long-run variance as $\widehat{\omega}_1^2 := \widehat{\gamma}_{0,1} + 2\sum_{j=1}^{T-2} k(j/M)\widehat{\gamma}_{j,1}$, with $\widehat{\gamma}_{j,1} := \frac{1}{T}\sum_{t=j+2}^T \widehat{u}_{t,1}\widehat{u}_{t-j,1}$. In addition to regularity conditions on u_t, consistency of $\widehat{\omega}_1^2$ depends upon the kernel function $k(\cdot)$ and the rate of divergence of M such that $M \to \infty$ and $M/T \to 0$ as $T \to \infty$ (for a discussion, see Jansson, 2002).

The coefficient and t-statistic based one-step PP unit root tests are given by

$$Z_{\alpha,1} := T(\widehat{\alpha}_1 - 1) - \frac{1}{2}\left(\widehat{\omega}_1^2 - \widehat{\sigma}_1^2\right)\left(T^{-2}\sum_{t=2}^T \widetilde{y}_{t-1,1}^2\right)^{-1}, \tag{8}$$

$$Z_{t,1} := \frac{\widehat{\sigma}_1}{\widehat{\omega}_1} t_{\alpha_1} - \frac{1}{2}\left(\widehat{\omega}_1^2 - \widehat{\sigma}_1^2\right)\left(\widehat{\omega}_1^2 T^{-2}\sum_{t=2}^T \widetilde{y}_{t-1,1}^2\right)^{-1/2}. \tag{9}$$

2.2. Two-Step Approach

The two-step detrending approach is very similar, yet there are some subtle differences that will matter. The two-step approach is based on first detrending the series y_t to then estimate a regression similar to (7) using the detrended data. Thus, we have $\widetilde{y}_{t,2} := y_t - D_t'\widehat{\theta}$, with $\widehat{\theta} := \left(\sum_{t=1}^T D_t D_t'\right)^{-1}\sum_{t=1}^T D_t y_t$ and $\widetilde{y}_{t-1,2} := y_{t-1} - D_{t-1}'\widehat{\theta}$. Consequently the OLS estimator $\widehat{\alpha}_2$ of α used in this approach is given by

$$\widehat{\alpha}_2 := \frac{\sum_{t=2}^T \widetilde{y}_{t,2}\widetilde{y}_{t-1,2}}{\sum_{t=2}^T \widetilde{y}_{t-1,2}^2},$$

and the corresponding two-step residuals, used to compute the two-step estimates of σ^2 and ω^2, are given by $\widehat{u}_{t,2} := \widetilde{y}_{t,2} - \widehat{\alpha}_2 \widetilde{y}_{t-1,2}$.

The two-step PP tests, $Z_{\alpha,2}$ and $Z_{t,2}$ say, are defined as before—with $\widetilde{y}_{t-1,2}$ and the two-step estimates $\widehat{\alpha}_2$, $\widehat{\sigma}_2^2$, and $\widehat{\omega}_2^2$ instead of the corresponding one-step quantities—in (8) and (9).

2.3. Asymptotic Results

It is well known that for deterministic polynomial trends the asymptotic distribution of $T(\widehat{\alpha}-1)$ is the same for both the one-step and two-step approaches. Thus, when one appeals to a consistency result for an estimator of ω^2, the asymptotic distributions of the PP tests are identical for the one- and two-step versions of the tests, and it holds that

$$Z_{\alpha,i} \Rightarrow -c + \frac{\int_0^1 \widetilde{V}_c(r)dW(r)}{\int_0^1 \widetilde{V}_c(r)^2 dr}, \qquad Z_{t,i} \Rightarrow -c\sqrt{\int_0^1 \widetilde{V}_c(r)^2 dr} + \frac{\int_0^1 \widetilde{V}_c(r)dW(r)}{\sqrt{\int_0^1 \widetilde{V}_c(r)^2 dr}},$$

for $i = 1, 2$.

Remark 1. Further differences between one- and two-step detrending occur in the case of more general deterministic components because then also the one- and two-step limits of $T(\widehat{\alpha}-1)$ may differ. In particular, the numerator of the two-step limit contains an additional term:

$$-\int_0^1 \widetilde{V}_c(r)\dot{D}(r)dr \left(\int_0^1 D(r)D(r)'dr\right)^{-1} \int_0^1 D(r)V_c(r)dr,$$

with $\dot{D}(r)$ denoting the (corresponding quantity in more general cases) first difference of $D(r)$ as defined before for the polynomial trend case. This term is, clearly, not zero in general, but in the case of polynomial trends it holds that $\int_0^1 \widetilde{V}_c(r)\dot{D}(r)dr = 0$, because in that case the span of $\dot{D}(r)$ is contained in the span of $D(r)$. For an example where this term is not zero, see Perron and Vogelsang (1992), who include a mean shift dummy in the deterministic component.

The fact that these limiting distributions rely upon the consistency of $\widehat{\omega}_i^2$ implies that the asymptotic distributions do not capture the influence of the randomness in $\widehat{\omega}_i^2$ on the resulting test statistics. In particular, the choices with respect to both the kernel function and the bandwidth are not reflected in the asymptotic distribution, yet affect the finite-sample performance of $\widehat{\omega}_i^2$ and thus of the PP test statistics.

This limitation of conventional asymptotic theory is addressed with fixed-b theory by means of deriving an asymptotic approximation for $\widehat{\omega}_i^2$ under the assumption that $M = bT$, where $b \in (0,1]$ is held fixed as $T \to \infty$. In practice for a given sample with T observations and a given value of M, one would use the fixed-b

limit corresponding to the value of $b = M/T$. The fixed-b limit of $\widehat{\omega}_i^2$ depends on the asymptotic behavior of the scaled partial sums of $\widehat{u}_{t,i}$, which is shown in Lemma 1 to differ between the one- and two-step approaches. This in turn implies that also the one- and two-step PP test statistics will have different limits when using the fixed-b approximation.

LEMMA 1. *Assume that the data are generated by (1) with the errors, u_t, fulfilling the FCLT (3). As $T \to \infty$ it holds for $0 \leq r \leq 1$ that*

$$T^{-1/2} \sum_{t=2}^{[rT]} \widehat{u}_{t,1} \Rightarrow \omega H_{1,c}(r),$$

$$H_{1,c}(r) := \widehat{W}(r) - \left(\frac{\omega^2 \int_0^1 \widetilde{V}_c(s) dW(s) + \lambda}{\omega^2 \int_0^1 \widetilde{V}_c(s)^2 dr} \right) \int_0^r \widetilde{V}_c(s) ds, \tag{10}$$

$$T^{-1/2} \sum_{t=2}^{[rT]} \widehat{u}_{t,2} \Rightarrow \omega H_{2,c}(r),$$

$$H_{2,c}(r) := \widetilde{\dot{V}}_c(r) - \left(\frac{\omega^2 \int_0^1 \widetilde{V}_c(s) dW(s) + \lambda}{\omega^2 \int_0^1 \widetilde{V}_c(s)^2 dr} \right) \int_0^r \widetilde{V}_c(s) ds. \tag{11}$$

Lemma 1 shows that, in the limit, the scaled residual partial sums have leading terms that differ between the two approaches. The leading terms reflect the impact of the detrending method, and the second terms are identical, because they essentially reflect the estimation of α, which is for the considered deterministic components asymptotically equivalent in both cases.

Based on the preceding partial sum results, the fixed-b limits of $\widehat{\omega}_i^2$, $i = 1, 2$, can be expressed in terms of the processes $H_{1,c}(r)$ and $H_{2,c}(r)$. As is common in fixed-b theory, the results depend upon b and the shape of the kernel function in ways outlined in Definition 1.

DEFINITION 1. *With $H(r)$ denoting a scalar stochastic process define the stochastic process $P(b, k, H)$ as follows:*

(i) If $k''(x)$ exists and is continuous, then

$$P(b, k, H) = -\frac{1}{b^2} \int_0^1 \int_0^1 k'' \left(\frac{r-s}{b} \right) H(r) H(s) dr ds + \frac{2}{b} H(1) \int_0^1 k' \left(\frac{1-r}{b} \right) H(r) dr + H(1)^2.$$

(ii) If $k(x)$ is continuous, $k(x) = 0$ for $|x| \geq 1$, and $k(x)$ is twice continuously differentiable everywhere except for possibly $|x| = 1$, then

$$
\begin{aligned}
P(b,k,H) = &-\frac{1}{b^2}\iint_{|r-s|\le b} k''\left(\frac{r-s}{b}\right)H(r)H(s)drds \\
&+\frac{2}{b}k'_{-}(1)\int_0^{1-b} H(r)H(r+b)dr \\
&+\frac{2}{b}H(1)\int_{1-b}^{1} k'\left(\frac{1-r}{b}\right)H(r)dr + H(1)^2,
\end{aligned}
$$

with $k'_{-}(1) = \lim_{h\to 0}(k(1)-k(1-h))/h$.

(iii) If $k(x) = 1-|x|$ *for* $|x| \le 1$ *and* $k(x) = 0$ *otherwise, then*

$$
\begin{aligned}
P(b,k,H) = &\frac{2}{b}\int_0^1 H(r)^2dr - \frac{2}{b}\int_0^{1-b} H(r)H(r+b)dr \\
&-\frac{2}{b}H(1)\int_{1-b}^{1} H(r)dr + H(1)^2.
\end{aligned}
$$

Using the quantities just defined the following proposition gives the fixed-b limits of $\widehat{\omega}_i^2$ and of the PP tests for both the one- and two-step approaches.

PROPOSITION 1. *Assume that the data are generated by (1) with the errors,* u_t*, fulfilling the FCLT (3). Furthermore assume that* $M = bT$*, with* $b \in (0,1]$ *fixed and the subscript* $i \in \{1,2\}$ *refers again to the one- and two-step detrending approaches. Then as* $T \to \infty$

$$\widehat{\omega}_i^2 \Rightarrow \omega^2 P(b,k,H_{i,c}),$$

with $P(b,k,H)$ *as given in Definition 1 and*

$$Z_{\alpha,i} \Rightarrow -c + \frac{\int_0^1 \widetilde{V}_c(r)dW(r) + \frac{1}{2}\left(1-P(b,k,H_{i,c})\right)}{\int_0^1 \widetilde{V}_c(r)^2dr},$$

$$Z_{t,i} \Rightarrow -c\sqrt{\frac{\int_0^1 \widetilde{V}_c(r)^2dr}{P(b,k,H_{i,c})}} + \frac{\int_0^1 \widetilde{V}_c(r)dW(r) + \frac{1}{2}\left(1-P(b,k,H_{i,c})\right)}{\sqrt{P(b,k,H_{i,c})\int_0^1 \widetilde{V}_c(r)^2dr}}.$$

The proposition shows that under fixed-b asymptotics the asymptotic null distributions of the PP tests exhibit certain distinct features. First, the limits are nonpivotal given the dependence on σ^2 and ω^2 via the dependence on $H_{i,c}(r)$. This asymptotic result indicates that the finite-sample performance of the PP tests will be sensitive to the serial correlation structure in u_t even for moderate to large sample sizes, which matches the well-known finite-sample problems of the PP statistics documented in the literature. Second, the fixed-b limits are different for the one- and two-step approaches, with these differences occurring via $\widehat{\omega}_i^2$. Third, as is common when using fixed-b asymptotics, the choices of bandwidth and kernel are captured by the asymptotic approximation as reflected by $P(b,k,H_{i,c})$. Notice that, when $P(b,k,H_{i,c}) = 1$, the standard PP asymptotic distributions are

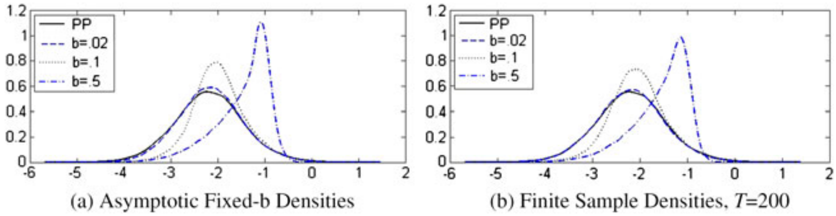


FIGURE 1. Densities for Z_t statistic, Bartlett kernel, two-step, intercept + trend, $\rho = 0$, $\varphi = 0$. PP denotes the standard PP limiting density.

obtained, which is exactly as expected because using a consistent estimator of ω^2 exactly coincides with $P(b, k, H_{i,c}) = 1$.

It is instructive to compare the fixed-b limiting distributions of the PP tests to the standard limiting distribution of the PP tests. Figure 1a plots asymptotic densities for the Z_t statistic using the Bartlett kernel in the intercept + trend model using two-step detrending for the case where there is no additional serial correlation in the model ($\lambda = 0$). The densities were computed using simulation methods. The simulations were performed using partial sums of 1,000 independent and identically distributed (i.i.d.) $N(0,1)$ random errors to approximate the Wiener process that drives the limits. The number of replications is 100,000. Asymptotic fixed-b densities are given for $b = 0.02, 0.1, 0.5$ along with the standard PP asymptotic density. When $b = 0.02$ there is a small difference between the standard PP limit and the fixed-b limit. As b increases, there is a greater discrepancy between the standard PP density and the fixed-b density. To gauge the relevance of these asymptotic results for finite-samples, we simulate the finite-sample densities of the PP statistic for the case of $T = 200$ using the following ARMA(1,1) model for u_t:

$$u_t = \rho u_{t-1} + \varepsilon_t + \varphi \varepsilon_{t-1}, \tag{12}$$

where ε_t is a sequence of i.i.d. $N(0,1)$ random variables for $t = 0, 1, ..., T$ and $u_0 = 0$. We again used 100,000 replications. Figure 1b plots finite-sample densities for the case of $\rho = 0$, $\varphi = 0$ using the same values of b as in Figure 1a. Also included in Figure 1b is the standard asymptotic PP density. When taken in isolation, Figure 1b indicates two things. First, the finite-sample density of the Z_t statistic is sensitive to the bandwidth. Second, unless the bandwidth is small, the standard asymptotic PP density is inadequate. If we compare Figures 1a and 1b, we see that the asymptotic fixed-b densities capture the impact of the bandwidth on the finite-sample behavior of Z_t quite well.

Figure 2a plots the asymptotic fixed-b densities for the case of $b = 0.02$ but where additional serial correlation is included in the model. We again use the ARMA(1,1) specification for u_t. As the ARMA(1,1) parameters change, we see that the fixed-b densities move with them. This happens because the $P(b, k, H_{i,c})$ process in the fixed-b limit changes as ρ and φ change. Figure 2b provides the

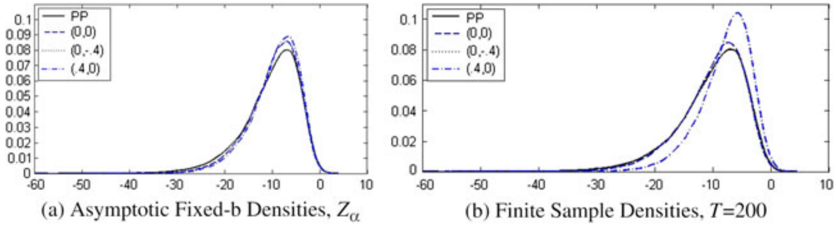


FIGURE 2. Densities for Z_α statistic, Bartlett kernel, two-step, intercept + trend, (ρ, φ) pairs, $b = 0.02$. PP denotes the standard PP limiting density.

analogous finite-sample densities for the case of $T = 200$. We again see that the patterns in the finite samples are captured by the fixed-b limits but not by the standard PP limit.

Because the fixed-b limiting random variables capture much of the finite-sample behavior of the long-run variance estimators used to construct the PP statistics and because the fixed-b limits depend on the kernel, bandwidth, and serial correlation nuisance parameters, the fixed-b theory provides a nice theoretical explanation for the sometimes poor finite-sample properties of the PP statistics when the traditional PP asymptotic distribution is used to generate critical values. Because the fixed-b limit is not pivotal, we cannot easily obtain critical values, and this motivates a simple modification of the PP statistics that does deliver a pivotal fixed-b limit result.

3. MODIFIED PP UNIT ROOT TESTS

The reason for the nonpivotal limits of the PP tests under fixed-b asymptotics is that the scaled partial sums of the residuals $\widehat{u}_{t,i}$ are not, as has been shown in Lemma 1, directly proportional to ω^2, because of the dependencies in the processes $H_{i,c}$. It is, however, straightforward to modify the residuals $\widehat{u}_{t,i}$ to obtain the needed asymptotic proportionality for a pivotal fixed-b limit result. For obtaining this result it is in fact sufficient to construct residuals using modified estimators of α. Define the modified estimators as

$$\widehat{\alpha}_i^m := \widehat{\alpha}_i + \frac{\frac{1}{2}\widehat{\sigma}_i^2}{T^{-1}\sum_{t=2}^T \widetilde{y}_{t-1,i}^2},$$

for $i \in \{1, 2\}$.

Using the preceding modification, the one-step modified residuals can be written as

$$\widehat{u}_{t,1}^m := \widetilde{y}_{t,1} - \widehat{\alpha}_1^m \widetilde{y}_{t-1,1} = \widetilde{y}_{t,1} - \left(\widehat{\alpha}_1 + \frac{\frac{1}{2}\widehat{\sigma}_1^2}{T^{-1}\sum_{t=2}^T \widetilde{y}_{t-1,1}^2}\right)\widetilde{y}_{t-1,1}$$
$$= \widetilde{u}_t - \left(\frac{T^{-1}\sum_{t=2}^T \widetilde{y}_{t-1,1}u_t + \frac{1}{2}\widehat{\sigma}_1^2}{T^{-2}\sum_{t=2}^T \widetilde{y}_{t-1,1}^2}\right) T^{-1}\widetilde{y}_{t-1,1}. \qquad \textbf{(13)}$$

For the two-step approach the modified residuals can be written as

$$\widehat{u}_{t,2}^m := \widetilde{y}_{t,2} - \widehat{\alpha}_2^m \widetilde{y}_{t-1,2} = \widetilde{y}_{t,2} - \left(\widehat{\alpha}_2 + \frac{\frac{1}{2}\widehat{\sigma}_2^2}{T^{-1}\sum_{t=2}^T \widetilde{y}_{t-1,2}^2}\right)\widetilde{y}_{t-1,2}$$

$$= u_t - (D_t - D_{t-1})'\left(\widehat{\theta} - \theta\right) - \left(\frac{T^{-1}\sum_{t=2}^T \widetilde{y}_{t-1,2}u_t + \frac{1}{2}\widehat{\sigma}_2^2}{T^{-2}\sum_{t=2}^T \widetilde{y}_{t-1,2}^2}\right) T^{-1}\widetilde{y}_{t-1,2}. \tag{14}$$

The following lemma gives the limit of the scaled partial sums of the modified residuals.

LEMMA 2. *Assume that the data are generated by (1) with the errors, u_t, fulfilling the FCLT (3). Consider the modified residuals, $\widehat{u}_{t,i}^m$, as given by (13) and (14) for the one- and two-step approaches. Then as $T \to \infty$ it holds that*

$$T^{-1/2}\sum_{t=2}^{[rT]} \widehat{u}_{t,1}^m \Rightarrow \omega H_{1,c}^m(r),$$

$$H_{1,c}^m(r) := \widehat{W}(r) - \left(\frac{\int_0^1 \widetilde{V}_c(s)dW(s) + \frac{1}{2}}{\int_0^1 \widetilde{V}_c(s)^2 dr}\right)\int_0^r \widetilde{V}_c(s)ds, \tag{15}$$

$$T^{-1/2}\sum_{t=2}^{[rT]} \widehat{u}_{t,2}^m \Rightarrow \omega H_{2,c}^m(r),$$

$$H_{2,c}^m(r) := \dot{\widetilde{V}}_c(r) - \left(\frac{\int_0^1 \widetilde{V}_c(s)dW(s) + \frac{1}{2}}{\int_0^1 \widetilde{V}_c(s)^2 dr}\right)\int_0^r \widetilde{V}_c(s)ds. \tag{16}$$

Thus, we see that the processes $H_{i,c}^m(r)$ are free of nuisance parameters and the limit of the scaled partial sums of the modified residuals is proportional to ω. These results now form the basis for modified PP tests that employ estimators of ω^2 using the modified residuals rather than the original ones. We denote the corresponding estimators of the long-run variance as $\widetilde{\omega}_i^2$ in what follows, and using $\widetilde{\omega}_i^2$ instead of $\widehat{\omega}_i^2$ defines the modified PP tests.

The following proposition provides the fixed-b limit distributions of the modified statistics.

PROPOSITION 2. *Assume that the data are generated by (1) with the errors, u_t, fulfilling the FCLT (3). Furthermore assume that $M = bT$, with $b \in (0,1]$ fixed and the subscript $i \in \{1,2\}$ refers again to the one- and two-step detrending approaches. Then as $T \to \infty$*

$$\widetilde{\omega}_i^2 \Rightarrow \omega^2 P\left(b, k, H_{i,c}^m\right)$$

and

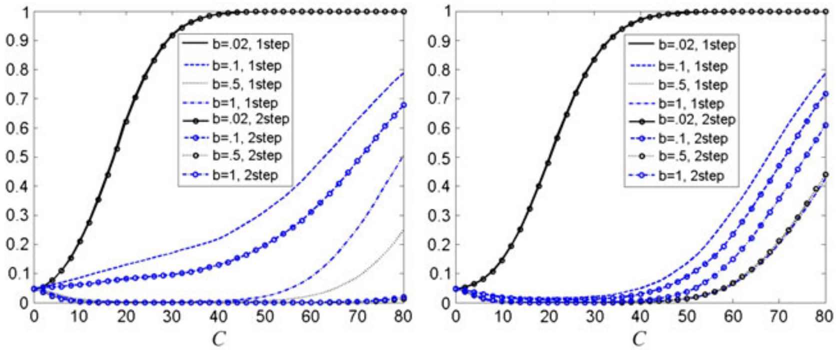
$$Z_{\alpha,i}^{m} \Rightarrow -c + \frac{\int_0^1 \widetilde{V}_c(r)dW(r) + \frac{1}{2}\left(1 - P\left(b,k,H_{i,c}^{m}\right)\right)}{\int_0^1 \widetilde{V}_c(r)^2 dr},$$

$$Z_{t,i}^{m} \Rightarrow -c\sqrt{\frac{\int_0^1 \widetilde{V}_c(r)^2 dr}{P\left(b,k,H_{i,c}^{m}\right)}} + \frac{\int_0^1 \widetilde{V}_c(r)dW(r) + \frac{1}{2}\left(1 - P\left(b,k,H_{i,c}^{m}\right)\right)}{\sqrt{P\left(b,k,H_{i,c}^{m}\right)\int_0^1 \widetilde{V}_c(r)^2 dr}}.$$

Because the processes $H_{i,c}^{m}(r)$ do not depend upon nuisance parameters, the processes $P\left(b,k,H_{i,c}^{m}\right)$ are also free of nuisance parameters, which leads to pivotal fixed-b limiting distributions of the modified PP statistics. Thus, under the null hypothesis of a unit root ($c = 0$), critical values can be simulated for given deterministic components, b and kernel function. These are, as already mentioned in the Introduction, available upon request for five kernels (Bartlett, Bohman, Daniell, Parzen, and QS) for the specifications without deterministic component, with intercept only, and with intercept and linear trend. For the latter two specifications of the deterministic component the fixed-b critical values differ between one- and two-step detrending. The values for b in these tables range from 0.02 to 1 with a mesh of size 0.02.

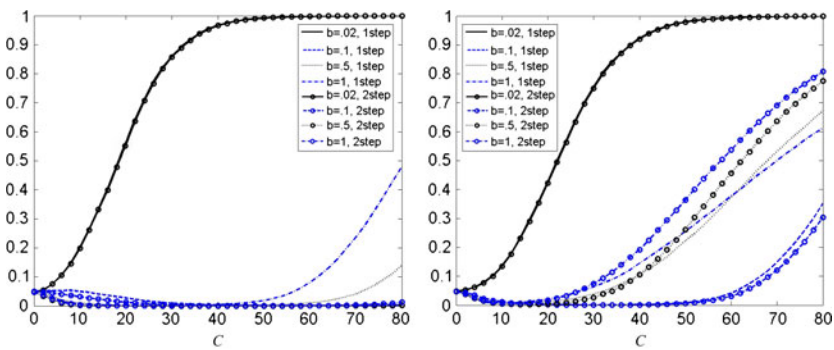
For nonzero values of c we can use the results of Proposition 2 to compute local asymptotic power (LAP) of the modified statistics. Because the limits in Proposition 2 depend on the kernel, bandwidth, and form of detrending, we can use LAP to make predictions about the impact of kernel, bandwidth, and detrending choices on finite-sample power. We simulate LAP for the mentioned five kernels and a selection of values of b using the same methods as used for Figures 1a and 2a now using 5,000 replications. LAP is computed for a grid over c running from 0 to 80 with steps of size 2. Rejections are computed using the $c = 0$ asymptotic critical value for a given kernel, bandwidth, detrending combination.

We report results for the Bartlett and QS kernels. The results for the other kernels are qualitatively similar. We report results for the intercept + trend model. Patterns are qualitatively similar for the intercept only model. Figures 3a and 3b plot LAP for the Bartlett kernel, whereas Figures 4a and 4b plot LAP for the QS kernel. The first notable pattern in these figures is the sensitivity of power to the choice of b. In many cases power decreases as b increases, but in other cases power is nonmonotonic in b. For example Z_t^m using the QS kernel has good power when $b = 0.02$, but power drops very quickly when b is increased to 0.1. Then, as b is increased further, power increases but stays well below power when $b = 0.02$. The second notable pattern is that, except for $b = 0.02$, there are clear differences in power between the one- and two-step detrending approaches with the greatest differences occurring for larger values of b. The third notable pattern is that power is often very low, often close to zero, when b is not small and c takes on small to medium values. The fourth notable pattern is that the kernel matters for power unless $b = 0.02$, in which case power is similar for both kernels in all cases.



(a) Z_α^m Statistic (b) Z_t^m Statistic

FIGURE 3. Local asymptotic power, Bartlett kernel, intercept + trend.



(a) Z_α^m Statistic (b) Z_t^m Statistic

FIGURE 4. Local asymptotic power, QS kernel, intercept + trend.

Finally, there are substantial differences in power between the Z_α^m and Z_t^m statistics except when $b = 0.02$. The local asymptotic power analysis suggests that small bandwidths are much preferable to nonsmall bandwidths. For this reason we report in Table 1 asymptotic fixed-b critical values of the modified statistics for the case of $b = 0.02$ for the Bartlett and QS kernels.

4. FINITE-SAMPLE BEHAVIOR

For the sake of brevity we only include a small selection of finite-sample results obtained by performing extensive simulations. In particular we only report some results for the sample size $T = 200$ for the Bartlett and QS kernels for the t-statistic tests. Qualitatively similar results are available also for the coefficient tests and sample size $T = 100$. The number of replications is 5,000 for each

TABLE 1. Critical values for the modified PP statistics Z_α^m and Z_t^m for $b = 0.02$

D_t	Detr.	Kernel	Statistic	90%	95%	97.5%	99%
Intercept	1-step	Bartlett	Z_α^m	−10.617	−13.070	−15.526	−18.690
			Z_t^m	−2.515	−2.780	−3.055	−3.334
		QS	Z_α^m	−10.492	−12.876	−15.268	−18.325
			Z_t^m	−2.503	−2.786	−3.038	−3.318
	2-step	Bartlett	Z_α^m	−10.660	−13.120	−15.593	−18.728
			Z_t^m	−2.516	−2.780	−3.055	−3.331
		QS	Z_α^m	−10.541	−12.938	−15.317	−18.356
			Z_t^m	−2.505	−2.787	−3.038	−3.315
Intercept + trend	1-step	Bartlett	Z_α^m	−16.723	−19.508	−22.082	−25.205
			Z_t^m	−3.024	−3.280	−3.498	−3.757
		QS	Z_α^m	−16.346	−19.090	−21.541	−24.533
			Z_t^m	−2.994	−3.248	−3.462	−3.720
	2-step	Bartlett	Z_α^m	−16.874	−19.670	−22.259	−25.369
			Z_t^m	−3.035	−3.292	−3.509	−3.769
		QS	Z_α^m	−16.545	−19.260	−21.734	−24.703
			Z_t^m	−3.009	−3.261	−3.477	−3.734

experiment. The selected results, for a narrow set of statistics and nuisance parameters, are meant to be illustrative in capturing the main observations in relation to the finite-sample predictions of the asymptotic theory. In particular, the fixed-b theory suggests that under the unit root null hypothesis: (i) the traditional PP unit root tests will be sensitive to nuisance parameters and the choice of kernel and bandwidth, (ii) the modified PP tests will be more robust to nuisance parameters and will be robust to the choice of kernel and bandwidth when the asymptotic fixed-b critical values are used, and (iii) there can be a difference between the one- and two-step detrending approaches.

The data generating process is given by (1) and (2) where we set $\theta = \mathbf{0}$ without loss of generality. We generate u_t according to the ARMA(1,1) model given by (12). The standard PP test uses the usual unit root asymptotic distribution critical values and is labeled PP. The modified PP test, labeled $PP(fb)$, uses the fixed-b asymptotic critical values corresponding to the limits given in Proposition 2 for $c = 0$.

In Figures 5–7 we display empirical null rejection probabilities at the 5% nominal level. The results are reported for a grid of bandwidths given by $M = 2, 4, 6, \ldots, 198, 200$, indexed by the corresponding value of $b = M/200$, with this grid corresponding to the grid for which fixed-b critical values have been simulated. In each figure, the rejections are reported for both one- and two-step detrending. Figures 5 and 6 give results for the case where u_t has no serial correlation ($\rho = 0$, $\varphi = 0$). Figures 5a and 6a give results for the intercept only case and Figures 5b and 6b give results for the intercept + trend case. Several patterns

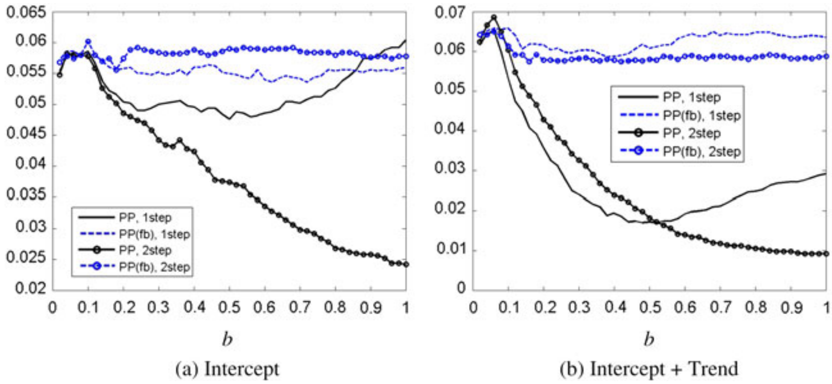


FIGURE 5. Empirical null rejections, Z_t^m statistic, Bartlett kernel, $\rho = 0$, $\varphi = 0$, $T = 200$.

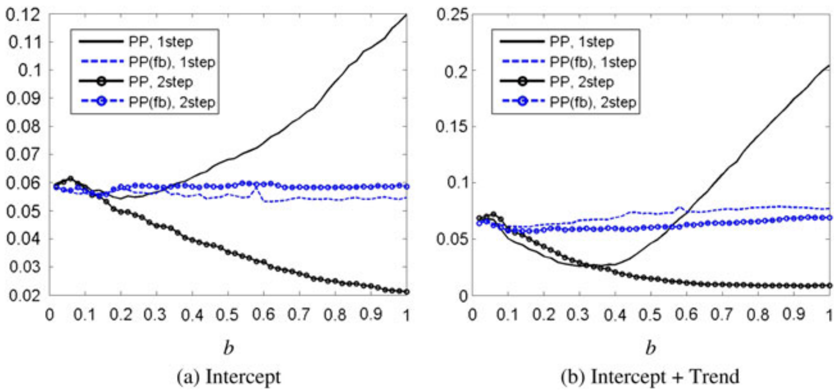


FIGURE 6. Empirical null rejections, Z_t^m statistic, QS kernel, $\rho = 0$, $\varphi = 0$, $T = 200$.

stand out in the figures. The rejections of the $PP(fb)$ tests are close to 0.05 regardless of the bandwidth, which indicates that the fixed-b critical values are doing an adequate job of capturing the dependence of the finite-sample distribution on the bandwidth. In contrast, unless a small bandwidth is used, the PP tests have rejections that are not close to 0.05, and the rejections show a sensitivity to the bandwidth and to the kernel. This is consistent with the predictions of Proposition 1. Comparing the one-step with the two-step approach, we see that for the PP tests, there are stark differences in rejection probabilities between the two approaches as predicted by Proposition 1. In contrast, the $PP(fb)$ tests have similar rejections for both one-step and two-step detrending, indicating that the critical values based on the one- and two-step fixed-b limits correctly capture the dependence upon detrending method.

Figures 7a and 7b give results for autoregressive and moving average errors for the Bartlett kernel in the intercept + trend case. Results are similar for the

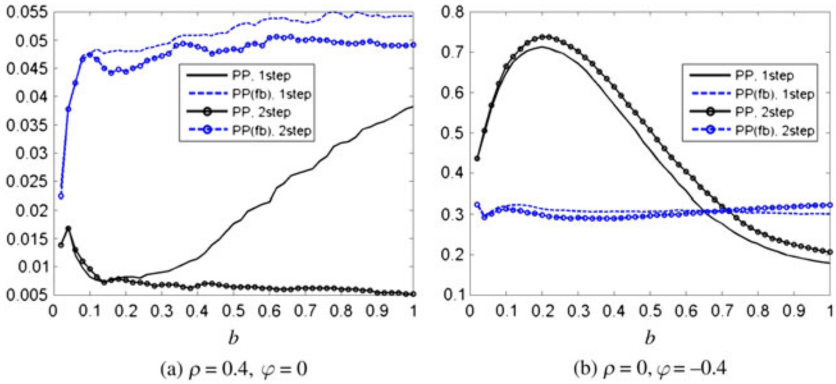
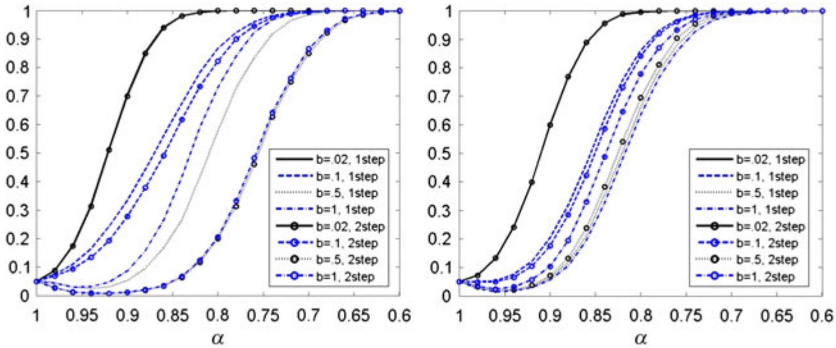


FIGURE 7. Empirical null rejections, Z_t^m statistic, Bartlett kernel, intercept + trend, $T = 200$.

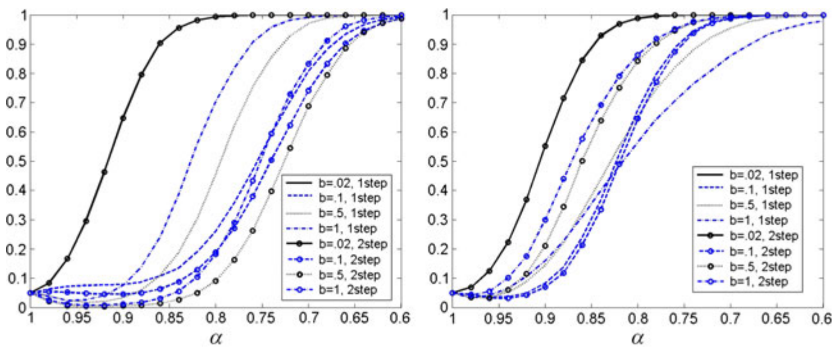
intercept only model and the QS kernel. Figure 7a has $\rho = 0.4, \varphi = 0$, and Figure 7b has $\rho = 0, \varphi = -0.4$. Rejections of the PP statistics are systematically different from 0.05 regardless of the bandwidth. This is consistent with the dependence on serial correlation nuisance parameters indicated by the fixed-b limits of the PP tests. There is a noticeable difference between the one-step and two-step approaches, especially in Figure 7a. The rejections of the PP statistics are even more distorted in Figure 7b, with substantial overrejections possible for some bandwidths. The rejections of the $PP(fb)$ statistics are very different. In Figure 7a rejections are close to 0.05 except when the bandwidth is small and underrejections occur. In contrast, when $\rho = 0, \varphi = -0.4$, rejections are inflated above 0.05 for the $PP(fb)$ statistics but are not sensitive to the bandwidth. This general tendency to overreject when there is a negative moving average component is well documented in the literature; see Perron and Ng (1996). In unreported simulations, we found that the overrejection problem becomes even more severe for $\varphi = -0.8$, as one would expect.

We now turn to some limited finite-sample power results to assess the adequacy of the LAP results for making predictions about the finite-sample power of the $PP(fb)$ statistics. We focus on the intercept + trend case given that results are similar for the intercept only case. We use the same values of b as for the LAP results. We only report results for $\rho = 0$, $\varphi = 0$ and power is size adjusted in all cases. Figures 8a and 8b depict power of the Z_α^m and Z_t^m statistics with the Bartlett kernel. Figures 9a and 9b give results for the QS kernel. The general patterns in Figures 8 and 9 are similar to the LAP results. Power is highest for $b = 0.02$, and power can be much lower for other values of b. There are noticeable differences in power between the one- and two-step approaches, and there are also clear differences in power between the Z_α^m and Z_t^m statistics. The one notable difference between the LAP power curves and finite-sample power curves is that power with $T = 200$ is generally higher than what is predicted by the LAP



(a) Z_t^m Statistic (b) Z_t^m Statistic

FIGURE 8. Finite-sample power (size adjusted), Bartlett kernel, Intercept + trend, $\rho = 0$, $\varphi = 0$, $T = 200$.



(a) Z_t^m Statistic (b) Z_t^m Statistic

FIGURE 9. Finite-sample power (size adjusted), QS kernel, intercept + trend, $\rho = 0$, $\varphi = 0$, $T = 200$.

analysis. The values of $\alpha = 0.9, 0.8, 0.7, 0.6$ with $T = 200$ correspond to values of $c = 20, 40, 60, 80$. Comparing Figure 4b with Figure 9b, notice that power with $c = 40$ is very low in Figure 4b (except for $b = 0.02$) whereas power with $\alpha = 0.8$ is much higher in Figure 9b. Although the LAP analysis adequately captures the general patterns of power with respect to dependency on bandwidth, model, one- and two-step detrending, etc., the accuracy of the magnitudes of power predicted by the LAP analysis is not so impressive.

Our final two figures compare finite-sample power of the PP and $PP(fb)$ statistics with several variants of the Perron and Ng (1996) modified statistics, referred to as PN statistics. We include four variants of these statistics, two of them based on nonparametric long-run variance estimation and two of them based

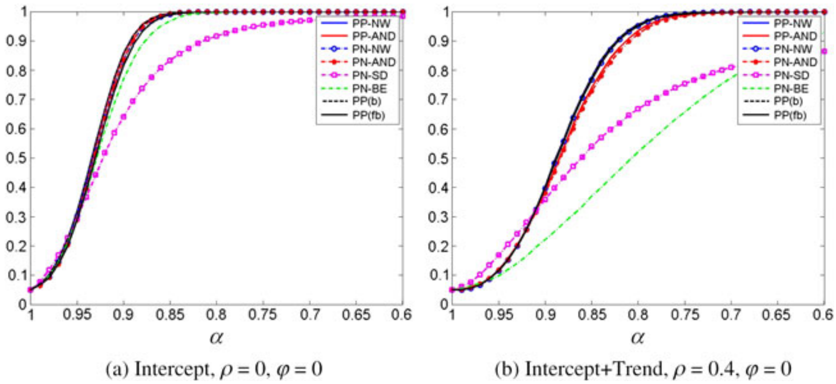


FIGURE 10. Finite-sample power (size adjusted), one-step, $T = 200$; PP(b),PP(fb) are Z_t, Z_t^m with Bartlett kernel and $b = 0.02$.

on estimating the long-run variance by fitting an autoregression. The nonparametric tests use the data dependent methods devised by Newey and West (1994) and Andrews (1991), labeled $PN - NW$ and $PN - AND$. The two variants of autoregressive model–based modified test statistics follow either Said and Dickey (1984) or Berk (1974) and are labeled $PN - SD$ and $PN - BE$. The results are given for the one-step approach and the t-statistic versions of the tests. The PP and $PP(fb)$ statistics use the Bartlett kernel with $b = 0.02$. We also implement variants of the PP statistic using the two data dependent nonparametric long-run variance estimation methods just mentioned to choose the bandwidths. These statistics are denoted $PP - NW$ and $PP - AND$ and for these and also for $PN - NW$ and $PN - AND$ the Bartlett kernel is used. Figure 10a gives results for the intercept only case with $\rho = 0$, $\varphi = 0$ whereas Figure 10b gives results for the intercept + trend case with $\rho = 0.4$, $\varphi = 0$. We see that the power of the PP and $PP(fb)$ statistics is similar whether PP uses $b = 0.02$ or a data dependent bandwidth. The power of the nonparametric PN statistics is similar, but the power of the two autoregression PN statistics tends to be lower. When a small bandwidth is used, the $PP(fb)$ statistics have power that is competitive with the original PP tests and other more size-robust tests.

5. SUMMARY AND CONCLUSIONS

The fixed-b theory developed in this paper provides an alternative theoretical explanation for the finite-sample dependence of the traditional PP unit root tests on serial correlation in the errors driving the unit root. Unlike the traditional consistency approximation for the long-run variance estimators used in the PP tests, the fixed-b theory also indicates a finite-sample difference between one-step and two-step detrending. Both local asymptotic power simulations and finite-sample simulations show that there can be large differences between the one-step and two-step detrending approaches in terms of both null rejection probabilities and power. We propose modified PP test statistics that have asymptotically pivotal fixed-b limits. The fixed-b limits depend on the kernel and bandwidth used for the long-run variance estimators, and the fixed-b limits are different for the one-step and two-step approaches. In finite samples, the modified PP statistics, when used with fixed-b critical values, have null rejections close to the nominal level unless the serial correlation in the innovations to the unit root process behaves similarly to an overdifferenced stationary process.

Although the modified PP tests are a clear improvement over the traditional PP tests, further modifications would be necessary to make them reliable in practice in terms of adequate size control in the presence of a negative moving average component. One approach that could be highly fruitful is to apply the wild bootstrap with recoloring to the modified PP tests. A recent paper by Cavaliere and Taylor (2009) has shown that size distortions of the traditional PP tests are substantially reduced when the wild bootstrap with recoloring is applied. See Table 3, panel (a), of Cavaliere and Taylor (2009). An interesting topic for future research is to study the behavior of a bootstrap version of the modified PP tests. A theoretical analysis of this approach would require a unification and potential extensions of bootstrap theory results given by Cavaliere and Taylor (2009) (unit root aspects of the bootstrap) and bootstrap theory results given by Gonçalves and Vogelsang (2011) (fixed-b aspects of the bootstrap).

REFERENCES

Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.

Berk, K.N. (1974) Consistent autoregressive spectral estimates. *Annals of Statistics* 2, 489–502.

Cavaliere, G. & A.M.R. Taylor (2009) Heteroskedastic time series with a unit root. *Econometric Theory* 25, 1228–1276.

Gonçalves, S. & T.J. Vogelsang (2011) Block bootstrap HAC robust tests: The sophistication of the naive bootstrap. *Econometric Theory* 27, 745–791.

Hashimzade, N. & T.J. Vogelsang (2008) Fixed-b asymptotic approximation of the sampling behavior of nonparametric spectral density estimators. *Journal of Time Series Analysis* 29, 142–162.

Jansson, M. (2002) Consistent covariance matrix estimation for linear processes. *Econometric Theory* 18, 1449–1459.

Kiefer, N.M. & T.J. Vogelsang (2005) A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory* 21, 1130–1164.

Newey, W. & K. West (1994) Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61, 631–653.

Perron, P. & S. Ng (1996) Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *Review of Economic Studies* 63, 435–463.

Perron, P. & T.J. Vogelsang (1992) Testing for a unit root in a time series with a changing mean: Corrections and extensions. *Journal of Business & Economic Statistics* 10, 467–470.

Phillips, P.C.B. (1987) Towards a unified asymptotic theory for autoregression. *Biometrika* 74, 535–547.

Phillips, P.C.B. & P. Perron (1988) Testing for a unit root in time series regression. *Biometrika* 75, 335–346.

Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.
Said, S.E. & D.A. Dickey (1984) Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 71, 599–607.
Schwert, W. (1989) Tests for unit roots: A Monte Carlo investigation. *Journal of Business & Economic Statistics* 7, 147–159.
Sims, C.A., J.H. Stock, & M.W. Watson (1990) Inference in linear time series models with some unit roots. *Econometrica* 58, 113–144.

APPENDIX: Proofs

Proof of Lemma 1. For both methods of detrending, the residuals can be written as

$$\widehat{u}_{t,i} = \widetilde{y}_{t,i} - \alpha \widetilde{y}_{t-1,i} - T(\widehat{\alpha}_i - \alpha)\frac{\widetilde{y}_{t-1}}{T},$$

with differences occurring asymptotically only in the first part $\widetilde{y}_{t,i} - \alpha\widetilde{y}_{t-1,i}$. We first consider one-step detrending, in which case we obtain

$$\widetilde{y}_{t,1} - \alpha\widetilde{y}_{t-1,1} = \widetilde{u}_t = u_t - D_t'(D_T'D_T)^{-1}D_T'U_T.$$

This implies that $T^{-1/2}\sum_{t=2}^{[rT]}\left(\widetilde{y}_{t,1} - \alpha\widetilde{y}_{t-1,1}\right) \Rightarrow \omega\widehat{W}(r)dr$. Under the assumptions stated it is well known that $T(\widehat{\alpha}_1 - \alpha) \Rightarrow \left(\left(\omega^2\int_0^1 \widetilde{V}_c(s)dW(s) + \lambda\right) / \left(\omega^2\int_0^1 \widetilde{V}_c(s)^2dr\right)\right)$ and that $T^{-3/2}\sum_{t=2}^{[rT]}\widetilde{y}_{t-1,1} \Rightarrow \omega\int_0^r \widetilde{V}_c(s)ds$. Combining these three results establishes (10).

Let us now look at two-step detrending. In this case we can write

$$\begin{aligned}
\widetilde{y}_{t,2} - \alpha\widetilde{y}_{t-1,2} &= y_t - D_t'\widehat{\theta} - \alpha\left(y_{t-1} - D_{t-1}'\widehat{\theta}\right) \\
&= y_t^0 + D_t'\theta - D_t'\widehat{\theta} - \alpha\left(y_{t-1}^0 + D_{t-1}'\theta - D_{t-1}'\widehat{\theta}\right) \\
&= y_t^0 - \alpha y_{t-1}^0 - D_t'\left(\widehat{\theta} - \theta\right) + \alpha D_{t-1}'\left(\widehat{\theta} - \theta\right) \\
&= u_t - D_t'\left(\widehat{\theta} - \theta\right) + \left(1 - cT^{-1}\right)D_{t-1}'\left(\widehat{\theta} - \theta\right) \\
&= u_t - (D_t - D_{t-1})'\left(\widehat{\theta} - \theta\right) - cT^{-1}D_{t-1}'\left(\widehat{\theta} - \theta\right) \\
&= u_t - (D_t - D_{t-1})'\left(D_T'D_T\right)^{-1}D_T'Y_T^0 - cT^{-1}D_{t-1}'\left(D_T'D_T\right)^{-1}D_T'Y_T^0
\end{aligned}$$

with $Y_T^0 := \left[y_2^0, \ldots, y_T^0\right]'$. Defining $G_D := \text{diag}\left(1, T, \ldots, T^q\right)$, straightforward calculations give the standard results

$$\begin{aligned}
T^{-1/2}G_D\left(D_T'D_T\right)^{-1}D_T'Y_T^0 &= \left(T^{-1}G_D^{-1}D_T'D_TG_D^{-1}\right)^{-1}T^{-3/2}G_D^{-1}D_T'Y_T^0 \\
&\Rightarrow \omega\left(\int_0^1 D(s)D(s)'ds\right)^{-1}\int_0^1 D(s)V_c(s)ds
\end{aligned}$$

and

$$T^{-1}\sum_{t=2}^{[rT]} D_{t-1}'G_D^{-1} \Rightarrow \int_0^r D(s)'ds.$$

Less standard, but straightforward, is the result

$$\sum_{t=2}^{[rT]} (D_t - D_{t-1})' G_D^{-1} \Rightarrow \int_0^r \dot{D}(s)' ds.$$

Using these limits and (3), it easily follows that $T^{-1/2} \sum_{t=2}^{[rT]} (\widetilde{y}_{t,2} - \alpha \widetilde{y}_{t-1,2}) \Rightarrow \omega \dot{\widetilde{V}}_c(r)$. Combining this result with the unchanged results (compared to the one-step detrending) for the other two components establishes (11). ■

Proof of Proposition 1. The results of the proposition follow from the asymptotic results for the partial sum processes of the residuals established in Lemma 1, using similar arguments as in Hashimzade and Vogelsang (2008). Once the fixed-b limits for $\widehat{\omega}_i^2$, $i \in \{1, 2\}$ are established, the fixed-b limit distributions of the test statistics follow from using these when calculating the limiting distributions of the test statistics, as given in (8) and (9), with these expressions referring to one-step detrending and the two-step detrending versions of the test statistics similarly defined. ■

Proof of Lemma 2. The proof of the lemma builds heavily on the proof of Lemma 1, with in fact only the term comprising $\widehat{\alpha}_i^m$ being different in the expressions for the modified residuals compared to the residuals previously considered. For both detrending approaches it holds that

$$\begin{aligned}
T(\widehat{\alpha}_i^m - \alpha) &= T(\widehat{\alpha}_i - \alpha) + \frac{\frac{1}{2}\widehat{\sigma}_i^2}{T^{-2}\sum_{t=2}^{T} \widetilde{y}_{t-1,i}^2} \\
&\Rightarrow \frac{\omega^2 \int_0^1 \widetilde{V}_c(r) dW(r) + \lambda}{\omega^2 \int_0^1 \widetilde{V}_c(r)^2 dr} + \frac{\frac{1}{2}\sigma^2}{\omega^2 \int_0^1 \widetilde{V}_c(r)^2 dr} \\
&= \frac{\int_0^1 \widetilde{V}_c(r) dW(r) + \frac{1}{2}}{\int_0^1 \widetilde{V}_c(r)^2 dr},
\end{aligned}$$

from which the results of the lemma follow, because all other terms are unchanged compared to Lemma 1. ■

Proof of Proposition 2. Using the result from Lemma 2 the limit of $Z_{\alpha,i}^*$ follows immediately:

$$\begin{aligned}
T(\widehat{\alpha}_i^m - 1) &= T(\widehat{\alpha}_i^m - \alpha + \alpha - 1) = T(\widehat{\alpha}_i^m - \alpha) + T(\alpha - 1) \\
&= T(\widehat{\alpha}_i^m - \alpha) - c \\
&\Rightarrow -c + \frac{\int_0^1 \widetilde{V}_c(r) dW(r) + \frac{1}{2}}{\int_0^1 \widetilde{V}_c(r)^2 dr}.
\end{aligned}$$

The other results follow, similar to the results of Proposition 1, from the asymptotic result for the partial sum processes of the modified residuals established in Lemma 2 now using the fixed-b limit for the modified long-run variance estimators $\widetilde{\omega}_i^2$ in place of $\widehat{\omega}_i^2$. ■